

Russian Patent No. 2093968 C1

---

Job No.: 228-118625

Ref.: Russian Pat. No. 2093968/PU020436 RU/JMN(Della)/Order no. 8241

Translated from Russian by the McElroy Translation Company

800-531-9977

[customerservice@mcelroytranslation.com](mailto:customerservice@mcelroytranslation.com)

FEDERAL SERVICE FOR INTELLECTUAL PROPERTY, PATENTS AND TRADE MARKS  
DESCRIPTION OF AN INVENTION FOR A PATENT OF THE RUSSIAN FEDERATION  
RUSSIAN PATENT NO. 2093968 C1

Int. Cl.<sup>6</sup>: H 04 N 7/12  
Status: as of 25.07.2008 – no longer in force  
Filing No.: Claim: 95112554/09  
Filing Date: 1995.08.02  
Publication Date: 1997.10.20  
List of Documents Cited in Support  
of the Claim: U.S. Patent No. 5010401, class H 04 N 7/12, 1991.

METHOD OF ENCODING-DECODING IMAGES AND AN APPARATUS FOR ITS  
IMPLEMENTATION

Patent Holder: Closed joint-stock company "Techno-TM"  
Applicant: Closed joint-stock company "Techno-TM"

[Abstract]

The invention relates to the field of digital processing of video image signals. The basic areas of application of this invention are connected with compressing data on narrow-band communications channels. This method uses the principle of adaptive vector quantization. Use of a conveyer principle of data processing in combination with partial parallelization of computing allows the method to achieve a time frame close to television distribution using data encoding and decoding. 2 examples, 1 illustration.

Description of the invention

The invention relates to the field of image signal processing and can be used, in particular, for repeated compression (encoding) of images in real time with subsequent regeneration (decoding) of the compressed image to its initial state with saving of the fine details and characteristics, and containing the significant data in the initial image.

The process of compressing images can be used during transmission of video images from space from an orbiting craft over narrow-band communications links in real time; during transmission of data (images) in computer communications networks; over existing narrow-band and telephone communications channels; for economical archiving of space, medical, judicial and other video data; for rapid transmission of large quantities of data between data servers and personal computer video terminals, as well as in a number of other scientific and technological applications.

There are a number of well-known methods of "compression-regeneration" of images. For example, adaptive differential pulse-code modulation and methods of encoding with prediction [1] provide a simple hardware implementation of a system for work in real time. This method allows for 4-5 fold compression, but the quality of the regenerated image depends directly (degrades) with the level of noise in the communications channels.

A method of compression, based on orthogonal transformation, such as Fourier, Adamar, or Karhunen transformations, or other transformations, is resistant to impulse and high-frequency interference in communications channels. However compression with good quality by these methods is achieved only for simple images and does not retain much fine detail.

To improve the quality of displayed images compressed using orthogonal transformation, block encoding is used [1]. In this case the whole image is not compressed immediately, and the divisions (blocks) contain a variable number of pixels. This method is widely used for encoding television images in commercial systems of video recording and television broadcasting, as well as in production of computer video and is the basis of the JPEG and MPEG standards used in many countries for image compression.

The JPEG standard is the most widely used method of compression and has the following operational characteristics:

- a) piem of the image in digital form and its buffering;
- b) discrete cosine transformation (D.C.T.), during which the image is divided into blocks and each block undergoes the above-mentioned operation. This operation requires large computing power, during which loss of information does not occur. D.C.T. exhibits a variety of Fourier transformation, and has a reverse transformation. The sense of this operation is in the transition from a spatial representation of the image to its spectral representation and back again. Clipping the highest-frequency elements of the displayed image, it is possible in this and in other directions depending on quantity to choose between the quality of the image (regenerated) and the degree of compression. In this way D.C.T. creates a matrix, presenting the initial image, in the aggregation of matrices of frequency coefficients. The high-frequency elements of these matrices can be coarsened or discarded without a fundamental loss of information. The time needed for calculation of each element of the D.C.T matrix depends heavily on its size, which JPEG recommends should be in the form of elemental matrices of  $8 \times 8$  elements;

c) quantization. After the D.C.T operation each block of the image undergoes a quantization operation, i.e., it is split off from the following block by rounding to the corresponding element of the special quantization matrix. Here a basic loss of information occurs on account of the rounding, during which the result of splitting off is rounded to the nearest whole number with an established binary series number. In the matrix received after the quantization, the area of high frequencies usually turns out to be filled with zeros;

d) then RLE [run-length encoding] is performed on the resulting matrices (replication encoding), during which the remaining frequency coefficients are encoded. The coefficient of each matrix is bypassed at this path in a zig-zag manner, and the null spacings are likewise encoded by a standard RLE program. Then the information from all blocks is formed into a single stream;

e) "entropy" encoding. The stream produced by the RLE undergoes a standard method of compression through Huffman or arithmetic encoding;

f) during decoding all the preceding steps are again performed in reverse order, i.e. Huffman decoding is performed, then decoding by an RLE program, then regeneration of the D.C.T matrices through reverse D.C.T. Finally, the image is taken from the regenerated blocks.

Use of the discussed method, depending on the complexity of the processed image, can achieve a compression coefficient of 10-60. However, with large coefficients of compression the regenerated image begins to show bleeding of the borders of groups (blocks) of pixels [1] causing the regenerated image to consist of a large number of lightly colored spots. This leads to, for example, difficulty in vision while viewing a video compressed by this method. The regenerated image loses its fine structure and therefore the use of this approach is difficult for compression of images containing a large number of details, for example, aerial photography.

An alternative method of block encoding using orthogonal transformation is compression of the image by adaptive vector quantization [4]. In this case the image is also split into blocks and the collection of pixels in the block is described by a vector. Adaptive vector quantization (A.V.Q.) is realized through encoding of the mutual characteristics of pixels in the encoded block, and therefore the regenerated image after such encoding preserves the fine details at relatively high compression (up to 16 times and higher).

The analogs examined above are sufficiently effective in their own fields. However none of them are adequate across a sufficiently broad range of practical applications. At present this range includes applications of image compression that require a combination of the advantages of all these methods, above all with respect to regeneration of images in real time over existing narrow-band communications channels.

a) Transmission of compressed images on communications channels. This field is particularly relevant at present, while applications of visual inspection from space are growing and it

is necessary to transmit large quantities of video information on space communications channels with minimal loss of detail at the pace of high-speed space video transmitters.

b) Local and global computer information networks. The connection between computers in such networks is often the existing narrow-band (including telephone) communications lines, and therefore the use of "compression-regeneration" systems facilitates transmission of large quantities of video data within such networks.

c) Computing complexes for image processing have a narrow place in transmission of video data from PC servers to graphics video terminals. Such transmissions are usually done over standard digital communications channels, which have rather narrow bandwidth. Therefore the use of efficient compression and regeneration of video data increases the speed of analysis of an image in such computer systems.

d) A separate problem is archiving large quantities of computer data in digital format. This can be a bank of space observation video data, a bank of medical photographs, judicial video data, or other data. Use of effective "compression-regeneration" for archived images can decrease the storage requirements of such computer systems by a factor 1.5-2.

The two following methods most closely satisfy all of the listed requirements, i.e. JPEG and AVQ. A comparison of these methods shows that JPEG has a sufficiently good quality of regeneration of images using a compression coefficient up to 15, but above a coefficient of 10, fine detail begins to be lost. AVQ provides similar results to JPEG, but preserves fine detail at a compression coefficient of about 10. At larger compression coefficients AVQ shows better results than JPEG, and therefore is used as a prototype. Avoidance of the basic drawback of AVQ, i.e. minimizing the effect of noise in the communications channel, can be achieved for image compression by using principles taken from Kohonen neural networks [5] and using them in problems of image compression.

It was decided to use a classical AVQ method and modify it in such a way that it preserved the topology of the encoded data similar to Kohonen neural networks. For this the following was necessary. Since, as for an apparatus for study of Kohonen neural networks, the initial location of clusters must be specially arranged. to the degree that a cluster order number increases, the value of the spatial coordinates of its center changes monotonically. For this the initial centers of the clusters were arranged in a single diagonal of an  $n$ -dimensional vector space, and the number of clusters varied along the diagonal monotonically. Further, during the above-described process of adaptation of the location of the clusters, the initial topology of their distribution is altered, [and] clusters with adjoining number can appear in various corners of the quantified vector space. In order for this not to occur, after liquidation of each empty cluster and its opening, near the largest of the remaining clusters, a renumbering of all clusters occurs in the following manner. The newly opened cluster is given a number next to the number of the cluster located at the partition, while the previous or next

number of the clusters (depending on whether the number of the liquidated cluster was greater or less than the distributed cluster) is moved to one so that with the existing numbers [sic] and the freed number of the liquidated cluster is given to one of the remaining non-empty clusters. As a result, a robust method of studying Kohonen neural networks was developed for compression [and] regeneration of images. During 16-fold compression, the regenerated image retains its structure and the fine details of the original, and the average quadratic error of regeneration of the image is almost the same as with a classical AVQ method for image processing on which this method was based. If a neural network is used for image compression, for which it was never before used, the quality of its output exceeds the quality of the output of an AVQ method.

In this way a classical AVQ method is based on preliminary construction of code tables (code books, code libraries, libraries of references) based on a process of learning (adaptation) of the compression of a concrete class of images. In this process the libraries used do not retain the topology of the encoded data, which leads to large errors during decoding, caused by small errors in the encoded data, occurring for example, during transmission over communications channels. The proposed process of learning retains the indicated topology (vectors with similar content have similar numbers in the code library) and passes through the following image. The initial image is split into blocks of pixels of the specified size  $K \times L$ , each of which is interpreted as a vector in a  $K \times L$ -dimensional space. With a method of adaptive clusterization it is essential to divide the given space for "t" so that the density of the location of clusters corresponds to the density of the distribution of probability of vectors in the indicated vector space. Before beginning the adaptive clusterization, the centers of the indicated "t" clusters are designated and placed in this vector space according to some ordering scheme. Subsequently, on the basis of an analysis of a sample of the learned data (p-dimensional vectors, i.e. all blocks-pixels of the encoded image), adaptation of the location of these centers to the statistical properties of the encoded data takes place. At the same time, for each of the produced vectors, the deviation from the centers of the indicated clusters is determined according to the criteria of the minimum of the Euclidian distance. For each of the clusters the new position of the center is calculated as the average for all vectors in the clusters, empty clusters are liquidated and new clusters with a given number are created with a new spatial location by means of splitting in two the largest of the remaining clusters. One of the halves is given the number of the source cluster, and another number is again produced (liquidated in the second place of the spatial vector). The center of the source cluster is taken as these two new clusters, and two small random p-dimensional opposite vectors are added to the initial coordinates (in this way, the centers of the recreated clusters are distributed in the space). An important point in this is the retention of the previous order of location of the clusters in the code library, because after creation of the new cluster, reordering takes place with the goal of preserving the existing order. Subsequently, the learned sample of vectors is again produced for clusterization and this process is repeated

iteratively until its criteria are completed, for example, until the point at which the change of coordinates of the centers of the clusters are, according to the criteria, less than an assigned threshold.

Another important distinguishing characteristic of the proposed method of compression is the use of pseudo-floating points during storage and the use of reference libraries. Studies have shown that during preparation of code libraries by any of the methods in use, the best results for quality of images regenerated after compression are achieved in those cases in which the value of the vectors of the library are computed and used with a format with a floating point. However this is very ineffective from a computing point of view with respect to encoding. In this invention we propose using a library prepared with a floating point during encoding. Then the received value of the vectors is multiplied by 64 and rounded to the nearest integer. During this operation, first, a two-byte number is used with 14 significant bits, which closely emulates the floating point operation. Second, two remaining bits (missing from the two bytes) are sufficient so that during calculation of the minimal Euclidian distance in the process of encoding the sum of the quadrant of the difference of the vectors (reference and code), it does not exceed two bytes, which allows for convenient programmatic and hardware realization of the algorithm. Third, during calculation only integer arithmetic is used. It is natural that during encoding of the values of the vectors of an image, it preliminarily moves to 6 bits.

During encoding using a method of determining the smallest Euclidian distance between the running vector and the reference to each encoded block of pixels of the image, the reference vector from the encoding table is placed in the "nearest" position, i.e. a 16-component initial vector is replaced by a single-component number of the reference vector. In this way 16-fold compression of the image is performed, while in fact it proceeds due to a reduction in the statistical video data surplus as a result of the corresponding "learning" of the reference code vector table.

After transmission over the communications channel or "archiving", the compressed image can be decoded. During this process, the meaning of each pixel of the decoded image is the address of an entry into the decoding table, in which is stored the corresponding values of the reference vector, just as in the coding table, i.e. the number of the reference vector is changed during decoding to a  $4 \times 4$  block of pixels.

It is worth noting that this method has little sensitivity to differently encoded images and images for which a coding table is used; therefore a once written coding table is applicable for compression of a wide class of images (for example, a class of space photographs, portraits, etc.). The construction of coding tables and vector reference tables is accomplished in stationary conditions using a PC or an active neural network. Then they are loaded into a fixed memory of the compression-regeneration processes and used during state operations. If needed, the class of coded images can be changed by changing the fixed memory, or it is possible to arrange for the changing

of many code tables using narrower classes of coded images (to increase the accuracy of regeneration).

In this way the proposed method avoids the drawbacks and significantly increases the functional possibilities of the classical AVQ method. This is achieved through the use of the principles of Kohonen neural networks for compression and based on these principles, the construction of a reference library that preserves the topology of the encoded data, which gives the proposed method robustness in relation to a wide range of images and allows achievement of the following technological results:

- a) in the encoding process, finding a means of determining the minimal Euclidian distance from the reference table for each block of the image from the block nearest to it;
- b) replacement of a multi-component initial vector of the image by a single-component reference vector;
- c) regeneration of the initial vector after transmission of the encoded image using the number of a reference vector, enabling the contents of a block to be called from a reference library;
- d) the method features higher speed of decoding in relation to the described analogs, which makes it useful for reproduction of moving images.

The figure shows a functional diagram of an encoding-decoding apparatus.

The encoding-decoding apparatus contains: data sampling register 1, first and second blocks of random access memory 2 and 3, first commutator 4, block of random access memory of references 5, subtraction block 6, quadrator 7, reference choice controller 8, first, second, third, fourth accumulators 9-12, first, second, third adders 13, 14, 15, decision block 16, output register 17, third, fourth blocks of random access memory 18, 19, second commutator 20, block for reference address generation 21, block for construction and calculation of a reference address 22, block of random access memory of the reference 23, exit register 24.

The encoding-decoding apparatus functions in the following manner. The initial image in the form of a digital sample of the initial data arrives at the first and second blocks of random access memory 2 and 3 in 4 strings for supply of information to the following blocks. Then the difference between the reference and running values of the signal sample of the image is formed in the subtraction block 6. Then in quadrator 7, a quadrate difference signal is formed in parallel from 4 lines (from 1 of 4 read-outs). The results of the computation are placed on 8 channels simultaneously for further processing. The results are added in adders 9, 10, 11, and 12. Since the dimensions of the coded block are  $4 \times 4$ , the receipt of the quadrate difference for the whole vector is accomplished in 2 passes. At the same time, the following 8 read-outs are calculated from the buffer. The eight received sums are transferred to the addition block, where they are added twice in pairs to produce a total sum. The result is stored in the decision block 16 and is used to choose a minimum value of the vector in the coding reference library. Then the procedure repeats for the next block of the encoded



image until all the required part of the reference library is processed and there are no more blocks to process. In this way the architecture is built on a conveyor principle using parallel processing within the nodes. The cycle of the conveyor structure determines the maximum time of processing of the data in one of the nodes of the conveyor. After determination of the closest vector from the reference library, encoding of a line of the image takes place. The coding block is implemented using the following elements: register for sampling the incoming data, commutator, control and access block of the series 1533 and 1564 microprocessors, blocks of MS6264-20NS Texas Instruments random access memories 1 and 2, subtraction block 6 and quadrator 7, adders 13, 14, 15 in microprocessor TMS2210, blocks of random access memory 5. The reference libraries are built in an MS62256-20NC microprocessor, control processor of type TMS320C10Nh or TMS3220C25FNh. The input register of the image decoding block provides a connection with a bus, from which issues data in the form of an encoded image. A line of the encoded image moves to the buffered memory of blocks 18 and 19, which provide continuous reception and transmission of data. Data from one of the RAM storage devices is directed through the commutator to the block for generation of the physical address of the reference in memory, where the number of the vector from the reference library is transformed into the beginning address of the location in RAM storage. The beginning address of the reference moves to the block for memory address construction and counter. After reading out one line of the reference block in the apparatus for address formation, the beginning address of the following vector arrives and reading out of the line of the reference block at the given address occurs and so on. Read-out of the information proceeds from the reference memory, appears in the input register and then a line of the initial image is formed.

A decoder can be implemented from the following elements: input register, commutator, block for memory address construction and counter, output register (based on series 1553, 1554, 55PT17), RAM device for references based on MS62256-20NC (block for formation of the image based on series 174 and 1118 microprocessors).

## References

1. R.A. Uintu. Image encoding by means of transformation, TIHER, 1972, vol. 60, N7, p. 6981.
2. R. Radcheev, R. Frazier. Means of image compression for work with a scanner MIR PK, N 4, 1992, pp. 53-45
3. Nasrebadi N. M. King R.A. Image Coding Using Vector Quantization A. Review JEEE Trans. on Cjmm 76 vol. 36(8), 1988, pp.81-93.
4. S. Kuhn. Matrix processors at the SVIS [possibly Middle Volga Institute of Construction] Per. Sangl. M. Mir, 1991, p. 672.

5. US Patent N 5010401, H04N 7/12, 1991. Apparatus for encoding-decoding using vector quantization.

6. PCT N WO-90-09079, H04N 7/12. Method and apparatus for quantization.

### Claims

1. A method of encoding-decoding an image that enables reception of a digital signal of an image, which enables dividing it into blocks containing in each block a series of pixels that describe vectors, constructing a quantized vector space, forming a library of references from the signal corresponding to the first frame of the image by means of adaptation of the compression of concrete blocks of the image, calculating analogs in the library of the references for each of the blocks of the image, forming a signal of differences between the running value of the signal and the corresponding value from the reference library, forming a signal of the square of differences between the value of the signal and the corresponding value from the reference library, adding the squares of differences for eight blocks of the signal and then adding by pairs the received sums of the squares of differences to receive a total sum, storing the received code of the image block, reading this block in buffered memory with the subsequent formation of the regenerated image, distinguished by the fact, that the closest block from the reference library, corresponding to each block of the image, is located by means of determination of the minimal Euclidian distance, using the received total sum, and during decoding changing the multi-component initial vector of the image to a single-component number of a reference vector.

2. An apparatus for encoding-decoding an image, containing an encoder consisting of a register for the initial data, then a collection of blocks of reference RAM, a subtraction block, squarer and first accumulator, as well as a control block for electing a reference, whose output is connected to the input of a block of RAM of references, a decoder consisting of an input selection register, subsequently connected to a block for generation of an address, a block for composing the address of the reference in memory and counter, a block of RAM for references and an output register, distinguished by the fact, that the encoder contains first and second blocks of RAM for four lines, a first commutator, second, third and fourth accumulators, first, second and third adders and a decision block, while the first and second outputs of the selection register for input data are connected to the inputs of the first and second blocks of RAM, whose outputs are connected in turn to the first and second inputs of the first commutator, whose output is connected to the second input of the subtraction block, the output of the squarer is connected to the inputs of the second, third and fourth accumulators, the outputs of the first and second accumulators are connected to the first and second inputs of the first and second accumulators are connected [sic] to the first and second inputs of the second adder, whose outputs are connected to the first and second inputs of the third adder, whose output is connected to the input of the decision block, whose first output is the output of the

encoder, the second output of the decision block is connected to the input of the control block for selection of a reference, the second input of the block of reference RAM is connected to the input of the apparatus, the decoder includes the third and fourth RAM blocks for four lines and a second commutator, while the first and second outputs of the selection register are connected to the inputs of the third and fourth blocks of RAM, whose outputs are connected in turn to the first and second inputs of the second commutator, the output of the second commutator is connected to the input of the block for generation of the reference address, and the output of the output register is the output of the apparatus.

---

#### NOTIFICATION OF A PATENT FOR AN INVENTION

---

Code of the change in legal status	MM4A – early termination of patents of the Russian Federation due to non-payment within the established tax period to maintain the patent in force
Notification published	2002.06.10
Number of the bulletin	16/2002

---

#### FIGURES

Figure 1



(19) RU (11) 2093968 (13) C1

(51) 6 H04N7/12

ФЕДЕРАЛЬНАЯ СЛУЖБА  
ПО ИНТЕЛЛЕКТУАЛЬНОЙ СОБСТВЕННОСТИ,  
ПАТЕНТАМ И ТОВАРНЫМ ЗНАКАМ

(12) ОПИСАНИЕ ИЗОБРЕТЕНИЯ К ПАТЕНТУ РОССИЙСКОЙ ФЕДЕРАЦИИ

Статус: по данным на 25.07.2008 - прекратил действие

(21) Заявка: 95112554/09

(22) Дата подачи заявки: 1995.08.02

(45) Опубликовано: 1997.10.20

(56) Список документов, цитированных в отчете о  
поиске: US, патент, 5010401, кл. Н 04 N 7/12,  
1991.(71) Заявитель(и): Закрытое акционерное  
общество "Техно-ТМ"(73) Патентообладатель(и): Закрытое  
акционерное общество "Техно-ТМ"

#### (54) СПОСОБ КОДИРОВАНИЯ-ДЕКОДИРОВАНИЯ ИЗОБРАЖЕНИЙ И УСТРОЙСТВО ДЛЯ ЕГО ОСУЩЕСТВЛЕНИЯ

Изобретение относится к области цифровой обработки видеосигналов изображений. Основные сферы применения данного изобретения связаны с сокращением объема информации по узкополосным каналам связи. Способ использует принцип адаптивного квантования векторов. Применение конвейерного принципа обработки поступающей информации в сочетании с частичным распараллеливанием вычислений позволяет достичь масштаба времени, близкого к телевизионной разверстке при выполнении операций кодирования и декодирования информации. 2 с. п. ф-лы, 1 ил.

#### ОПИСАНИЕ ИЗОБРЕТЕНИЯ

Изобретение относится к области обработки сигналов изображений и может быть использовано, в частности, для многократного сжатия (кодирования) изображения в реальном масштабе времени с последующим восстановлением (декодированием) сжатого изображения до исходного состояния с сохранением его мелких деталей, особенностей и другой, содержащейся в исходном изображении значимой информации.

Процесс сжатия изображений может быть использован при передаче аэрокосмической видеоинформации с борта летательного аппарата на Землю по узкополосным каналам связи в реальном масштабе времени; при передаче информации (изображений) в сетях компьютерной связи; по существующим узкополосным и телефонным каналам связи; экономичной архивации аэрокосмической, медицинской, криминалистической и прочей видеоинформации; скоростной передаче больших массивов данных между накопителями информации и видеотерминалами ЭВМ, а также в ряде других приложений науки и техники.

Известен ряд способов "сжатия-восстановления" изображений. Так например, адаптивная дифференциальная кодоимпульсная модуляция и методы кодирования с предсказанием [1] обеспечивают простую аппаратную реализацию системы для работы в реальном масштабе времени. Этот метод обеспечивает сжатие в 4-5 раз, но качество восстановления изображений существенно зависит (снижается) от уровня шумов в каналах связи.

Методы сжатия, основанные на ортогональных преобразованиях таких, как преобразования Фурье, Адамара, Каррунена [1] и др. устойчивы к импульсным и высокочастотным помехам каналам связи. Однако сжатие с хорошим качеством посредством этих методов достигается только для простых изображений и не содержит большого количества мелких деталей.

Для повышения качества восстанавливаемых изображений, сжатых с применением ортогональных

преобразований, используется блочное кодирование [1] В этом случае сжатию подвергается не все изображение сразу, а его части, на которые оно разбивается (блоки), содержащие определенное (в частности, равное) количество пикселей. Такой способ широко используется при кодировании телевизионных изображений в коммерческих системах видеозаписи и телевизионного вещания, а также при создании компьютерных видеофильмов и положен в основу применяемых во многих странах стандартов сжатия изображений JPEG, MPEG.

Стандарт JPEG является наиболее распространенным методом сжатия и представляет собой следующую последовательность операций:

а/ прием изображений в цифровой форме и его буферизация;

б/ дискретное косинусное преобразование (Д.К.П.), при котором изображение разбивается на блоки и каждый блок подвергается упомянутой операции. Это операция требует большой вычислительной мощности, причем потери информации при этом не происходит. Д.К.П. представляет собой разновидность преобразования Фурье, имеет обратное преобразование. Смысл этой операции заключается в переходе от пространственного представления изображений к его спектральному представлению и наоборот. Отсекая наиболее высокочастотные элементы преобразованного изображения, можно в ту или иную сторону в зависимости от количества выбирать между качеством изображения (восстановленного) и степенью сжатия. Таким образом, Д.К.П. преобразует матрицу, представляющую исходное изображение, в совокупности матриц частотных коэффициентов. Высокочастотные элементы этих матриц могут быть огрублены или отброшены без особых потерь информации. Время, необходимое для вычисления каждого элемента матрицы Д.К.П. сильно зависит от ее размера, который JPEG рекомендует в виде элементарных матриц размером (8x8) элементов;

в/ квантование. После операции Д.К.П. каждый блок изображения подвергается операции квантования, т.е. делится с последующим округлением на соответствующий элемент специальной матрицы квантования. Здесь происходит основная потеря информации за счет округления, при котором результат деления округляется до ближайшего целого с установленным числом двоичных разрядов. Матрица, полученная после квантования, в области высоких частот обычно оказывается заполненной нулями;

г/ далее производится RLE-кодирование полученных матриц (кодирование повторов), при котором кодируются оставшиеся частотные коэффициенты. Коэффициенты каждой матрицы обходятся при этом зигзагообразно, нулевые промежутки кодируются также посредством стандартной программы RLE. Далее информация со всех блоков формируется в единый поток;

д/ кодирование "энтропии". Поток, полученный при RLE-кодировании подвергается стандартному методу сжатия по Хаффмену или арифметическому кодированию;

е/ при декодировании все происходит в строго обратном порядке, т.е. после декодирования по Хаффмену следует декодирование по программе RLE, затем восстановление матрицы Д.К.П. посредством обратного Д.К.П. наконец, сборка изображения из восстановленных блоков.

Применение рассмотренного метода в зависимости от сложности обрабатываемого изображения позволяет достичь коэффициент сжатия от 10 до 60. Однако при больших коэффициентах сжатия на восстановленном изображении начинают проступать границы групп (блоков) пикселей [1] а само по себе восстановленное изображение будет при этом состоять из большого числа легко различимых однотонных пятен, что делает, например, просмотр видеофильмов с использованием данного метода сжатия утомительным для зрения. Происходит потеря мелкой структуры обрабатываемого изображения, поэтому оказывается затруднительным использовать этот подход для сжатия изображений, содержащих большое количество деталей, например, аэрофотоснимков.

Альтернативой методам блочного кодирования с использованием ортогонального преобразования является сжатие изображений посредством адаптивного квантования векторов [4] В этом случае изображение также делится на блоки и набор пикселей в блоке описывается вектором. Адаптивное квантование векторов (А. К. В. ) обеспечивает кодирование взаимных характеристик пикселей в кодирующем блоке, поэтому восстановленное изображение после такого кодирования сохраняет мелкие детали при достаточно большом сжатии (до 16 раз и более).

Рассмотренные выше аналоги достаточно эффективны в своих областях. Однако ни один из них не обеспечивает работы в достаточно широкой сфере практических приложений. К такой сфере в настоящее время следует отнести такие области применения сжатия изображения, где необходимо сочетание преимуществ всех перечисленных методов и, прежде всего, в отношении восстановленных

изображений в реальном масштабе времени по существующим узкополосным каналам связи.

а/ Передача сжатых изображений по каналам связи. Эта область становится особенно актуальной в последние годы, когда расширяется применение космических средств визуального наблюдения и возникает необходимость передавать большие массивы видеoinформации по каналам космической связи с ограниченной пропускной способностью в темпе работы высокоскоростных космических видеодатчиков.

б/ Локальные и глобальные компьютерные информационные сети. Связь между компьютерами внутри таких сетей часто реализуется по существенным узкополосным (в том числе, и телефонным) линиям связи, и поэтому применение систем "сжатие-восстановление" облегчает передачу больших массивов видеоданных внутри таких сетей.

в/ Вычислительные комплексы обработки изображений имеют узким местом передачу видеоданных от накопителей ЭВМ к графическим видеотерминалам. Такие передачи обычно осуществляются по стандартным цифровым каналам связи, имеющим достаточно узкую полосу пропускания. Поэтому применение эффективного сжатия и восстановления видеоданных повышает скорость анализа изображений в таких вычислительных системах.

г/ Отдельную проблему представляет собой архивация больших массивов вычислительных данных, представленных в цифровой форме. Это может быть банк космических разведывательных видеоданных, банк медицинских снимков, криминалистических видеоданных и др. Применение эффективного "сжатия-восстановления" хранимых изображений позволяет на 1,5-2 порядка сократить требуемые размеры ресурсов памяти таких вычислительных систем.

Наиболее близко всем перечисленным требованиям удовлетворяют два последних рассмотренных метода, т.е. JPEG и АКВ. Приведенное сравнение этих методов показало, что JPEG имеет достаточно хорошее качество восстановленных изображений при коэффициенте сжатия до 15, но после 10 начинается исчезновение мелких деталей. Метод АКВ дает результаты, сопоставимые с JPEG, но при коэффициентах сжатия порядка 10 сохраняет мелкие детали. При больших коэффициентах сжатия метод АКВ показывает результаты лучшие, чем у JPEG, поэтому он и принимается за прототип. Отстройка от основного недостатка метода АКВ, т.е. устранение влияния шумов в канале связи, может быть осуществлена использованием для сжатия изображений принципов, заимствованных из нейронных сетей Кохонена [5] при использовании их в задачах сжатия изображений.

Было решено использовать классический метод АКВ и модифицировать его так, чтобы он сохранял топологию кодированных данных подобно нейронным сетям Кохонена. Для реализации этого необходимо следующее. Так же, как для устойчивого обучения нейронных сетей Кохонена, исходное положение кластеров должно быть специальным образом упорядочено так, чтобы по мере нарастания порядкового номера кластера значения пространственных координат его центра изменялись монотонно. Для этого исходные центры кластеров были расставлены по одной диагональной  $n$ -мерного векторного пространства, а номера кластеров менялись вдоль диагонали монотонно. Далее, в ходе описанного выше процесса адаптации положения кластеров исходная топология их расположения искажается, кластеры с соседними номерами могут оказываться в разных углах квантуемого векторного пространства. Чтобы этого не происходило, после ликвидации каждого пустого кластера и его открытия вблизи наибольшего из оставшихся производится переназначение номеров всех кластеров следующим образом. Вновь открываемому кластеру дается номер, соседний с номером кластера, подлежащего разбиению, при этом предыдущий или последующие номера кластеров (в зависимости от того, был ли номер ликвидированного кластера больше или меньше распределяемого) сдвигаются на единицу, чтобы с номерами уже существующих и чтобы освободившийся номер ликвидированного кластера был приписан одному из оставшихся непустых. В результате были выработаны устойчивые методы обучения нейронных сетей Кохонена для сжатия-восстановления изображений. При 16-кратном сжатии восстановленное изображение сохраняет структуру и мелкие детали исходного, а средне квадратичная ошибка восстановления изображения почти такая же, как у классического метода АКВ при обработке изображений, на которых этот метод был обучен. Если же нейронная сеть используется для сжатия изображений, которые ей ранее не предъявлялись, качество ее работы превосходит качество работы по методу А.К.В.

Таким образом, классический метод А.К.В. базируется на предварительном построении кодовых таблиц (кодовых книг, кодовых библиотек, библиотек эталонов) на основе процесса обучения (адаптации) сжатию конкретных классов изображений. При этом используемые библиотеки не сохраняют топологию кодируемых данных, что приводит к большим ошибкам при декодировании, вызываемым малыми ошибками в кодируемых данных, возникающих, например, при передаче по

каналам связи. Предлагаемый процесс обучения сохраняет указанную топологию (близкие по содержанию векторы имеют близкие значения номеров в кодовой библиотеке) и протекает следующим образом. Исходное изображение разбивается на блоки пикселей определенного размера  $K \times L$ , каждый из которых интерпретируется как вектор в  $n \times K \times L$ -мерном пространстве. Необходимо методом адаптивной кластеризации разделить данное пространство на  $m$  кластеров таким образом, чтобы плотности расположения кластеров соответствовала плотности распределения вероятности векторов в указанном векторном пространстве. Перед началом адаптивной кластеризации назначаются центры указанных  $m$  кластеров и некоторым упорядоченным образом размещаются в этом векторном пространстве. Затем на основании анализа выборки обучающих данных ( $n$ -мерных векторов, т.е. всех блоков-пикселей кодируемого изображения) производится адаптация положения этих центров к статистическим свойствам кодируемых данных. При этом для каждого из предъявленных векторов определяется отклонение от центров упомянутых кластеров по критерию минимума евклидова расстояния. Для каждого из кластеров новое положение центра вычисляется как среднее по всем попавшим в кластер векторам, пустые кластеры ликвидируются и новые кластеры с данным номером создаются с новым пространственным положением путем разбиения надвое самого большого из оставшихся кластеров. Одной из половин при этом приписывается номер материнского кластера, а другой номер вновь создаваемого (ликвидированного в другом месте векторного пространства). В качестве этих двух новых кластеров берется центр материнского, к исходным координатам которого прибавлены два небольших случайных  $n$ -мерных противоположно направленных вектора (таким образом, разносятся в пространстве центры вновь создаваемых кластеров). Важным моментом при этом оказывается сохранение принятого ранее порядка расположения кластеров в получаемой кодовой библиотеке, поэтому после образования нового кластера происходит их переупорядочивание с целью сохранения установленного порядка. После этого обучающая выборка векторов вновь предъявляется для кластеризации и этот процесс повторяется итеративно до выполнения критерия его завершения, например, до момента, когда изменение координат центров кластеров не станет по модулю меньше заданного порога.

Другой важной отличительной особенностью предлагаемого способа сжатия является использование режима псевдоплавающей точки при хранении и применении библиотек эталонов. Дело в том, как показали исследования, что при подготовке кодовых библиотек любым из применяемых способов наилучшие результаты по качеству восстановленных после сжатия изображений достигаются в том случае, когда значения векторов библиотеки рассчитываются и используются в формате с плавающей точкой. Однако это очень неэффективно с вычислительной точки зрения при реализации процедуры кодирования. В данном изобретении предлагается использовать при кодировании библиотеку, подготовленную в режиме с плавающей точкой. Затем полученные значения векторов умножаются на 64 и округляются до ближайшего целого. При этом, во-первых, используются двухбайтовые числа с 14 значащими разрядами, которые хорошо эмулируют режим работы с плавающей точкой. Во-вторых, двух оставшихся разрядов (недостающих до двух байт) достаточно, чтобы при расчете минимального евклидова расстояния в процессе кодирования сумма квадрата разностей векторов (эталонного и кодируемого) не превышала двух байт, что обеспечивает удобную программную и аппаратную реализацию алгоритма. В-третьих, при вычислениях используется только целочисленная арифметика. Естественно, что при выполнении кодирования значения векторов изображения сдвигаются предварительно на 6 разрядов.

В процессе кодирования путем определения наименьшего евклидова расстояния между текущим вектором и эталонным каждому кодирующему блоку пикселей изображения ставится в соответствие "ближайший к нему" эталонный вектор из таблицы кодирования, т.е. 16-компонентный входной вектор заменяется однокомпонентным номером эталонного вектора. Таким образом, осуществляется 16-кратное сжатие изображения, причем фактически оно происходит за счет сокращения статистической избыточности видеоданных вследствие соответствующего "обучения" таблицы эталонных кодовых векторов.

После передачи по каналу связи или "архивации" сжатое изображение может быть декодировано. При этом значение каждого пикселя декодируемого изображения является адресом входа в таблицу декодирования, в которой хранятся соответствующие значения эталонных векторов, такие же, как в таблице кодирования, т.е. номер эталонного вектора заменяется при декодировании блоком  $4 \times 4$  пиксела.

Следует отметить, что данный способ имеет малую чувствительность к несоответствию кодируемых изображений и изображений, на которые обучались таблицы кодирования, поэтому записанные однажды таблицы кодирования пригодны для сжатия широкого класса изображений (например, класс космических снимков, портреты и др.). Формирование таблиц кодирования и таблиц эталонных векторов осуществляется в стационарных условиях с применением ЭВМ или действующей

нейросети. Затем они закладываются в ПЗУ процессов сжатия-восстановления и используются при штатной эксплуатации. При необходимости может быть произведена замена класса кодируемых изображений путем замены ПЗУ, либо возможна организация применения многих таблиц кодирования при использовании более узких классов кодируемых изображений (для повышения точности восстановления).

Таким образом, предлагаемый способ устраняет недостатки и существенно расширяет функциональные возможности классического метода АКВ. Это достигается использованием для сжатия принципом нейронных сетей Кохонена и созданием на их основе библиотеки эталонов, сохраняющей топологию кодируемых данных, что делает предлагаемый метод устойчивым по отношению к широкому классу изображений и позволяет достичь следующих технических результатов:

а/ нахождение в процессе кодирования посредством определения минимального евклидова расстояния для каждого блока изображения ближайшего к нему из библиотеки эталонов;

б/ замена многокомпонентного входного вектора изображения однокомпонентным номером эталонного вектора.

в/ Восстановление исходного вектора после передачи закодированного изображения по номеру эталонного вектора, содержание блока которого вызывается из библиотеки эталонов;

г/ способ обладает более высокой скоростью декодирования по отношению к описанным аналогам, что делает его удобным при воспроизведении движущихся изображений.

На чертеже представлена функциональная схема устройства кодирования-декодирования.

Устройство кодирования-декодирования содержит: регистр выбора данных 1, первый и второй блоки оперативной памяти 2 и 3, первый коммутатор 4, блок оперативной памяти эталонов 5, блок вычитания 6, квадрататор 7, блок управления выбором эталона 8, первый, второй, третий, четвертый накопители 9-12, первый, второй, третий сумматоры 13, 14, 15, решающий блок 16, регистр выработки 17, третий, четвертый блоки оперативной памяти 18, 19, второй коммутатор 20, блок генерации адреса эталонов 21, блок формирования и счета адреса эталона 22, блок оперативной памяти эталона 23, выходной регистр 24.

Устройство кодирования-декодирования функционирует следующим образом. Входное изображение в виде цифровой выборки входных данных поступает в первый и второй блоки оперативной памяти 2 и 3 на 4 строки для обеспечения информацией последующих блоков. Далее в блоке вычитания 6 формируется разность эталонного и текущего значений выборки сигнала изображения. Далее в квадрататоре 7 формируется сигнал квадрата разности параллельно с 4-х строк (с 1 по 4 отсчеты). Результаты вычислений на дальнейшую обработку поступают по 8 каналам одновременно. Результаты накапливаются в накопителях 9, 10, 11, 12. Так как размер кодируемого блока составляет 4x4, то получение квадрата разности для всего вектора осуществляется за 2 прохода. При этом считываются следующие 8 отсчетов из буфера. Полученные восемь сумм передаются в блок суммирования, где они дважды попарно складываются для получения общей суммы. Результат запоминается в решающем блоке 16 и используется для выбора минимального значения вектора в библиотеке эталонов кодирования. Далее процедура повторяется для данного блока кодируемого изображения до тех пор, пока не будет обработана требуемая часть библиотеки эталонов и не будет выбран код данного блока изображения. Таким образом, архитектура строится по конвейерному принципу с применением внутри узлов распараллеливания обработки. Такт конвейерной структуры определяется наибольшим временем обработки данных в одном из узлов конвейера. После определения наиболее близкого вектора из библиотеки эталонов формируется закодированная строка изображения. Блок кодирования реализуется с использованием следующих элементов: регистр выбора входных данных, коммутаторы, блок управления и выборки микросхемы серий 1533, 1564, блоки оперативной памяти 1 и 2 микросхемы типа MS6264-20NS фирмы Texas Instruments, блок вычитания 6 и квадрататор 7, сумматоры 13, 14, 15 на микросхемах TMS2210, блоки оперативной памяти 5. Библиотеки эталонов строятся на микросхемах MS62256-20NC, управляющий процессор типа TMS320C10Nh или TMS3220C25FNh. Входной регистр блока декодирования изображений обеспечивает связь с шиной, откуда поступает информация в виде закодированного изображения. Строка закодированного изображения поступает в буферную память блоков 18 и 19, которые обеспечивают постоянный прием и выдачу данных. Через коммутатор данные из одного из ОЗУ направляются в блок генерации физического адреса эталона в памяти, где номер вектора из библиотеки эталонов преобразуется в начальный адрес расположения в пространстве ОЗУ. Начальный адрес эталона поступает в блок формирования и счетчик адреса памяти. После



считывания одной строки эталонного блока в устройство формирования адреса поступает начальный адрес следующего вектора и происходит считывание строки эталонного блока по данному адресу и т.д. Считанная информация направляется из памяти эталонов, поступает на входной регистр и далее на формирование строки выходного изображения.

Декодер может быть реализован на следующих элементах: входной регистр, коммутатор, блок формирования и счетчик адреса эталонной памяти, выходной регистр (выполняется на основании микросхем серий 1553, 1554, 55PT17), ОЗУ эталона на основе MS62256-20NC (блок формирования изображений на основе микросхем серий 174 и 1118).

#### Литература

1. Уинту Р.А. Кодирование изображений посредством преобразований, ТИИЭР, 1972 г. т.60, N7, с. 69-81.
2. Радчеев Р. Фразер Р. Средство сжатия изображений для работы со сканером МИР ПК, N 4, 1992 г. с.53-45.
3. Nasrebadi N. M. King R.A. Image Coding Using Vector Quantization A. Review JEEE Trans. on Cjmm 76 vol. 36(8), 1988, pp.81-93.
4. Кун С. Матричные процессоры на СВИС. Пер. с англ. М. Мир, 1991 г. с. 672.
5. Патент США N 5010401, НО4N 7/12, 1991. Устройство кодирования-декодирования с использованием векторного квантования.
6. PCT N WO-90-09079, НО4N 7/12. Метод и аппаратура для квантования.

#### ФОРМУЛА ИЗОБРЕТЕНИЯ

1. Способ кодирования-декодирования изображения, состоящий в том, что осуществляют прием цифрового сигнала изображения, разделяют его на блоки, содержащие набор пикселей в каждом блоке, которые описывают вектором, создают квантуемое векторное пространство, формируют из сигнала, соответствующего первому кадру изображения, библиотеку эталонов путем адаптации сжатию конкретных блоков изображения, в библиотеке вычисляют аналоги эталонов для каждого из блоков изображения, формируют сигнал разности между текущим значением сигнала и соответствующим ему значением из библиотеки эталонов, формируют сигнал квадрата разности между текущим значением сигнала и соответствующим ему значением из библиотеки эталонов, суммируют квадраты разностей для восьми блоков сигнала, полученные суммы квадратов разностей попарно складывают для получения общей суммы, запоминают полученный код блока изображения, при декодировании запомненный код блока изображения определяют адрес блока в библиотеке эталонов, считывают этот блок в буферную память с последующим формированием восстановленного изображения, отличающийся тем, что ближайший блок из библиотеки эталонов, соответствующий каждому блоку изображения, находят путем определения минимального евклидова расстояния, используя полученную общую сумму, при декодировании заменяют многокомпонентный входной вектор изображения однокомпонентным номером эталонного вектора.

2. Устройство для кодирования-декодирования изображения, содержащее кодер, включающий регистр выбора входных данных, последовательно соединенные блок оперативной памяти эталонов, блок вычитания, квадратор и первый накопитель, а также блок управления выбором эталона, выход которого подключен к входу блока оперативной памяти эталонов, декодер, содержащий входной регистр выборки, последовательно соединенные блок генерации адреса, блок формирования и счета адреса памяти эталона, блок оперативной памяти эталонов и выходной регистр, отличающееся тем, что в кодере введены первый и второй блоки оперативной памяти на четыре строки, первый коммутатор, второй, третий и четвертый накопители, первый, второй и третий сумматоры и решающий блок, при этом первый и второй выходы регистра выбора входных данных подключены к входам первого и второго блоков оперативной памяти, выходы которых соединены соответственно с первым и вторым входами первого коммутатора, выход которого подключен к второму входу блока вычитания, выход квадратора соединен с входами второго, третьего и четвертого накопителей, выходы первого и второго накопителей подключены к первому и второму входам первого и второго накопителей, подключены к первому и второму входам первого сумматора, а третьего и четвертого накопителей к первому и второму входам второго сумматора, выходы которых соединены с первым и вторым входами третьего сумматора, выход которого подключен к входу решающего блока, первый

выход которого есть выход кодера, второй выход решающего блока соединен с входом блока управления выбором эталона, второй вход блока оперативной памяти эталонов соединен с входом устройства, в декодер введены третий и четвертый блоки оперативной памяти на четыре строки и второй коммутатор, причем первый и второй выходы регистра выборки соединены с входами третьего и четвертого блоков оперативной памяти, выходы которых подключены соответственно к первому и второму входам второго коммутатора, выход второго коммутатора подключен к входу блока генерации адреса эталонов, выход выходного регистра есть выход устройства.

---

#### ИЗВЕЩЕНИЯ К ПАТЕНТУ НА ИЗОБРЕТЕНИЕ

---

Код изменения правового статуса	ММ4А - Досрочное прекращение действия патентов РФ из-за неуплаты в установленный срок пошлин за поддержание патента в силе
Извещение опубликовано	2002.06.10
Номер бюллетеня	16/2002

---

#### РИСУНКИ

Рисунок 1